Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	0	(identif$5 and features and duplicate and document and scores and reciprocal and fragments and words).clm.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2005/07/07 10:12

clm 09/713,733

Interference Search History.

12 Phrasier: a system for interactive document retrieval using keyphrases
Steve Jones, Mark S. Staveley
August 1999 **Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval**
Full text available: pdf(625.73 KB)   Additional Information: full citation, references, citings, index terms

Keywords: evaluation, interactive retrieval interface, keyphrase-based retrieval, query interface

13 Identifying topics by position
Chin-Yew Lin, Eduard Hovy
March 1997 **Proceedings of the fifth conference on Applied natural language processing**
Full text available: pdf(996.73 KB)   Additional Information: full citation, abstract, references, citings
Publisher Site

This paper addresses the problem of identifying likely topics of texts by their position in the text. It describes the automated training and evaluation of an Optimal Position Policy, a method of locating the likely positions of topic-bearing sentences based on genre-specific regularities of discourse structure. This method can be used in applications such as information retrieval, routing, and text summarization.

14 Using suffix arrays to compute term frequency and document frequency for all substrings in a corpus
Mikio Yamamoto, Kenneth W. Church
March 2001 **Computational Linguistics**, Volume 27 Issue 1
Full text available: pdf(1.80 MB)   Additional Information: full citation, abstract, references, citings
Publisher Site

Bigrams and trigrams are commonly used in statistical natural language processing; this paper will describe techniques for working with much longer n-grams. Suffix arrays (Manber and Myers 1990) were first introduced to compute the frequency and location of a substring (n-gram) in a sequence (corpus) of length N. To compute frequencies over all $N(N + 1)/2$ substrings in a corpus, the substrings are grouped into a manageable number of equivalence classes. In this way, a ...

15 How reliable are the results of large-scale information retrieval experiments?
Justin Zobel
August 1998 **Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval**
Full text available: pdf(1.16 MB)   Additional Information: full citation, references, citings, index terms

16 Document filtering with inference networks
Jamie Callan
August 1996 **Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval**
Full text available: pdf(809.41 KB)   Additional Information: full citation, references, citings, index terms

17

Search: ◉ The ACM Digital Library   ○ The Guide

+identifying, +duplicate, +documents, +scores

USPTO

☞ Feedback Report a problem Satisfaction survey

Published before November 2001   **Found 201 of 119,420**
Terms used <u>identifying</u> <u>duplicate</u> <u>documents</u> <u>scores</u>

Sort results by   [relevance ▼]   ◆ <u>Save results to a Binder</u>   Try an <u>Advanced Search</u>
                                                                            Try this search in <u>The ACM Guide</u>
Display results   [expanded form ▼]   ⁂ <u>Search Tips</u>
                                        ☐ Open results in a new window

Results 1 - 20 of 200   Result page: **1** <u>2</u> <u>3</u> <u>4</u> <u>5</u> <u>6</u> <u>7</u> <u>8</u> <u>9</u> <u>10</u>   <u>next</u>
Best 200 shown   Relevance scale ☐▭▬▨▩

1 <u>Experiments on incorporating syntactic processing of user queries into a document retrieval strategy</u>   ▩
   A. F. Smeaton, C. J. van Rijsbergen
   May 1988 **Proceedings of the 11th annual international ACM SIGIR conference on Research and development in information retrieval**

   Full text available: 📄 <u>pdf(1.94 MB)</u>   Additional Information: <u>full citation</u>, <u>abstract</u>, <u>references</u>, <u>citings</u>, <u>index terms</u>

   Traditional information has relied on the extensive use of statistical parameters in the implementation of retrieval strategies. This paper sets out to investigate whether linguistic processes can be used as part of a document retrieval strategy. This is done by predefining a level of syntactic analysis of user queries only, to be used as part of the retrieval process. A large series of experiments on an experimental test collection are reported which use a parser for noun phrases as part o ...

2 <u>Mechanized Indexing Methods and Their Testing</u>   ▬
   John O'Connor
   October 1964 **Journal of the ACM (JACM)**, Volume 11 Issue 4

   Full text available: 📄 <u>pdf(1.04 MB)</u>   Additional Information: <u>full citation</u>, <u>references</u>, <u>citings</u>, <u>index terms</u>

3 <u>Creating and evaluating multi-document sentence extract summaries</u>   ▬
   Jade Goldstein, Vibhu Mittal, Jaime Carbonell, Jamie Callan
   November 2000 **Proceedings of the ninth international conference on Information and knowledge management**

   Full text available: 📄 <u>pdf(186.89 KB)</u>   Additional Information: <u>full citation</u>, <u>references</u>, <u>citings</u>, <u>index terms</u>

4 <u>Efficient text summarization using lexical chains</u>   ▬
   H. Gregory Silber, Kathleen F. McCoy
   January 2000 **Proceedings of the 5th international conference on Intelligent user interfaces**

   Full text available: 📄 <u>pdf(586.39 KB)</u>   Additional Information: <u>full citation</u>, <u>abstract</u>, <u>references</u>, <u>citings</u>, <u>index terms</u>